

WHAT IS CLAIMED IS:

1. A method for computer cluster virtualization comprises:

5           selecting a distributed application;  
          retrieving a policy associated with the  
distributed application;  
          dynamically selecting one of a plurality of  
nodes;  
          resetting a boot image of the selected node  
10 based, at least in part, on the retrieved policy, the  
boot image compatible with the distributed application;  
and  
          associating a virtual disk image with the  
selected node based, at least in part, on the retrieved  
15 policy; and  
          executing at least a portion of the distributed  
application on the reset node using the associated  
virtual disk image.

20       2. The method of Claim 1, the application  
executing on a subset of the plurality of nodes and the  
method further comprising:

          comparing the subset of nodes with the  
retrieved policy; and  
25       selecting one of a plurality of compatible boot  
images based on the comparison.

3. The method of Claim 2, wherein comparing the subset of nodes with the retrieved policy comprises:

determining a count of nodes in the subset; and  
selecting the boot image based on a link in the  
5 policy and the count of nodes.

4. The method of Claim 2, each of the subset of nodes associated with one of the plurality of compatible boot images.

10

5. The method of Claim 1, wherein dynamically selecting one of the plurality of nodes comprises:

determining if one or more of the plurality of nodes is unutilized by a second distributed application;  
15 and

in response to at least one of the nodes being unutilized, selecting one of the unutilized nodes.

6. The method of Claim 5, in response to none of  
20 the nodes being unutilized, further comprising selecting one of the nodes utilized by the second distributed application based on one or more of the following:

the retrieved policy;  
low utilization of the selected node;  
25 priority of the selected distributed application; and  
compatibility of the selected node with the selected distributed application.

7. The method of Claim 6, wherein resetting the boot image of the selected node comprises:

automatically shutting down the selected node;  
resetting the boot image of the selected node;

5 and

restarting the selected node using the reset boot image.

8. The method of Claim 7, further comprising  
10 terminating any processes associated with the second distributed application prior to shutting down the node.

9. The method of Claim 1, the policy comprising a plurality of links to boot images, each link associated  
15 with one of a count of nodes compatible with the distributed application.

10. The method of Claim 1, the policy comprising one or more parameters for determining the timing of the  
20 selection of the node.

11. The method of Claim 1, the boot image comprising a remote boot image stored in a Storage Area Network (SAN).

12. The method of Claim 1, the node associated with a first boot image prior to the reset and associated with a second boot image from the reset, the first and second boot image differing in at least one of the following  
5 characteristics:

operating system;  
system configuration; and  
distributed application parameters.

10 13. The method of Claim 1, further comprising:  
determining that one of the plurality of nodes failed, the failed node executing at least a portion of the selected distributed application; and  
wherein selecting one of the plurality of nodes  
15 comprises selecting one of the remaining nodes in response to the failure.

14. The method of Claim 1, each of the plurality of nodes comprising the same processor architecture.  
20

15. The method of Claim 1, wherein selecting one of the plurality of nodes comprises selecting one of the plurality of nodes at a predetermined time.

16. Software for computer cluster virtualization operable to:

select a distributed application;  
retrieve a policy associated with the  
5 distributed application;  
dynamically select one of a plurality of nodes;  
reset a boot image of the selected node based,  
at least in part, on the retrieved policy, the boot image  
compatible with the distributed application; and  
10 associate a virtual disk image with the  
selected node based, at least in part, on the retrieved  
policy; and  
execute at least a portion of the distributed  
application on the reset node using the associated  
15 virtual disk image.

17. The software of Claim 16, the application  
executing on a subset of the plurality of nodes and the  
software further operable to:

20 compare the subset of nodes with the retrieved  
policy; and  
select one of a plurality of compatible boot  
images based on the comparison.

25 18. The software of Claim 17, wherein the software  
operable to compare the subset of nodes with the  
retrieved policy comprises software operable to:

determine a count of nodes in the subset; and  
select the boot image based on a link in the  
30 policy and the count of nodes.

19. The software of Claim 17, each of the subset of nodes associated with one of the plurality of compatible boot images.

5        20. The software of Claim 16, wherein the software operable to dynamically select one of the plurality of nodes comprises software operable to:

             determine if one or more of the plurality of nodes is unutilized by a second distributed application;  
10        and

             in response to at least one of the nodes being unutilized, select one of the unutilized nodes.

21. The software of Claim 20, in response to none  
15        of the nodes being unutilized, further operable to select one of the nodes utilized by the second distributed application based on one or more of the following:

             the retrieved policy;  
             low utilization of the selected node;  
20        priority     of     the     selected     distributed  
             application; and

             compatibility of the selected node with the selected distributed application.

22. The software of Claim 21, wherein the software operable to reset the boot image of the selected node comprises software operable to:

5                    automatically shut down the selected node;  
                    reset the boot image of the selected node; and  
                    restart the selected node using the reset boot  
image.

23. The software of Claim 22, further operable to  
10 terminate any processes associated with the second  
distributed application prior to shutting down the node.

24. The software of Claim 16, the policy comprising  
a plurality of links to boot images, each link associated  
15 with one of a count of nodes compatible with the  
distributed application.

25. The software of Claim 16, the policy comprising  
one or more parameters for determining the timing of the  
20 selection of the node.

26. The software of Claim 16, the boot image  
comprising a remote boot image stored in a Storage Area  
Network (SAN).

27. The software of Claim 16, the node associated with a first boot image prior to the reset and associated with a second boot image from the reset, the first and second boot image differing in at least one of the  
5 following characteristics:

operating system;  
system configuration; and  
distributed application parameters.

10 28. The software of Claim 16, further operable to:  
determine that one of the plurality of nodes failed, the failed node executing at least a portion of the selected distributed application; and

wherein the software operable to select one of  
15 the plurality of nodes comprises software operable to select one of the remaining nodes in response to the failure.

29. The software of Claim 16, each of the plurality  
20 of nodes comprising the same processor architecture.

30. The software of Claim 16, wherein the software operable to select one of the plurality of nodes comprises software operable to select one of the  
25 plurality of nodes at a predetermined time.



31. A system for computer cluster virtualization comprises:

a plurality of computing nodes; and

a management node operable to:

- 5           select a distributed application;
- retrieve a policy associated with the distributed application;
- dynamically select one of a plurality of nodes;
- reset a boot image of the selected node based,
- 10   at least in part, on the retrieved policy, the boot image compatible with the distributed application; and
- associate a virtual disk image with the selected node based, at least in part, on the retrieved policy; and
- 15           execute at least a portion of the distributed application on the reset node using the associated virtual disk image.

32. The system of Claim 31, the application  
20   executing on a subset of the plurality of nodes and the management node further operable to:

compare the subset of nodes with the retrieved policy; and

select one of a plurality of compatible boot images  
25   based on the comparison.

33. The system of Claim 32, wherein the management node operable to compare the subset of nodes with the retrieved policy comprises the management node operable to:

- 5       determine a count of nodes in the subset; and  
      select the boot image based on a link in the policy and the count of nodes.

34. The system of Claim 32, each of the subset of  
10 nodes associated with one of the plurality of compatible boot images.

35. The system of Claim 31, wherein the management node operable to dynamically select one of the plurality  
15 of nodes comprises the management node operable to:

- determine if one or more of the plurality of nodes is unutilized by a second distributed application; and  
      in response to at least one of the nodes being unutilized, select one of the unutilized nodes.

20

36. The system of Claim 35, in response to none of the nodes being unutilized, selecting one of the nodes utilized by the second distributed application based on one or more of the following:

- 25       the retrieved policy;  
      low utilization of the selected node;  
      priority of the selected distributed application;  
      and  
      compatibility of the selected node with the selected  
30 distributed application.

37. The system of Claim 36, wherein the management node operable to reset the boot image of the selected node comprises the management node operable to:

5        automatically shut down the selected node;  
      reset the boot image of the selected node; and  
      restart the selected node using the reset boot image.

38. The system of Claim 37, the management node  
10 further operable to terminate any processes associated with the second distributed application prior to shutting down the node.

39. The system of Claim 31, the policy comprising a  
15 plurality of links to boot images, each link associated with one of a count of nodes compatible with the distributed application.

40. The system of Claim 31, the policy comprising  
20 one or more parameters for determining the timing of the selection of the node.

41. The system of Claim 31, the boot image comprising a remote boot image stored in a Storage Area  
25 Network (SAN).

42. The system of Claim 31, the node associated with a first boot image prior to the reset and associated with a second boot image from the reset, the first and second boot image differing in at least one of the  
5 following characteristics:

operating system;  
system configuration; and  
distributed application parameters.

10 43. The system of Claim 31, the management node further operable to:

determine that one of the plurality of nodes failed, the failed node executing at least a portion of the selected distributed application; and

15 wherein the management node operable to select one of the plurality of nodes comprises the management node operable to select one of the remaining nodes in response to the failure.

20 44. The system of Claim 31, each of the plurality of nodes comprising the same processor architecture.

45. The system of Claim 31, wherein the management node operable to select one of the plurality of nodes  
25 comprises the management node operable to select one of the plurality of nodes at a predetermined time.